

## MOBILITY DATA BEYOND FACEBOOK & DATA SEEKING BEHAVIORS

Interviews and contact with 70+ people from research teams and government offices revealed that an overwhelming number of teams worked with Facebook mobility data rather than the many other mobility data sources used for COVID-19 response activities.

Apple - Mobility Trends Report
Camber - COVID-19 Mobility Data Network
Facebook - COVID-19 Mobility Data Network
Google - COVID-19 Community Mobility Reports
Safegraph - COVID-19 Data Consortium
Unicast - social distancing scoreboard
Cuebiq - mobility insights

While the User Feedback Project was biased toward Facebook data this brief focuses on the circumstances and feedback from individuals and teams about why they used other data sources. It presents a behavioral view of how teams navigated these challenges moving from purpose to meaning (see [Envisioning use of mobility data in practice](#)).

### MOBILITY DATA OVERLOAD OPTING FOR FACEBOOK DATA

Many interviewees in government organizations who were interviewed in this project shared that their teams and sometimes their leadership were assessing and exploring other types of mobility data. Some teams decided to use Facebook data after looking at other data sources. For example, the governor of Massachusetts had looked at Apple mobility trends reports but directed an analyst who was the data stream lead to focus on Facebook data as it was more granular than what they found available at the time from Apple Maps. One Chief Intelligence Officer (CIO) in California was noted to be scanning different mobility sources as well.

*While the User Feedback Project was biased toward Facebook data this brief focuses on the circumstances and feedback from individuals and teams about why they used other data sources.*

This project was not able to interview governors, CIOs, and other key decision-makers so a deeper understanding of the drivers for seeking mobility data, and their direct sentiments on why they preferred one versus the other are unknown.

## WHEN FACEBOOK DATA FAILS - OPTING FOR AN ALTERNATIVE

For some collaborations in the CMDN, using Facebook mobility data was not a good match, and this prompted the research-practitioner teams to explore other datasets together. For example, Rebecca Kahn, a Postdoctoral Research Fellow at Harvard University, worked with Melissa Schigoda, the Director of the Office of Performance and Accountability for the City of New Orleans, and found that the Facebook data analysis was less useful due to tile sparsity. They had hoped to use the mobility data as part of a dashboard for the mayor, to help target messaging to neighborhood groups. They found that the lack of data in each tile along with unavailable data from the early months of the pandemic significantly limited its use. This was partly due to the fact that Facebook protocols only begin to collect mobility data from a specific region at the start time of engagement. Not all teams joined the Network at the onset of the pandemic.

Rebecca and Melissa continued to adapt and subsequently chose Camber data. The analysis presented results by census tract which was much more useful for the team. Melissa's team also used traffic data and Safegraph location specific data (e.g., visits per location) which was useful for planning business reopenings as well as assisting with health department personal protective equipment (PPE) distributions. Teams in Tennessee, New Orleans, and Philadelphia faced obstacles in geographic specificity, for which some of these challenges are further described in the Brief ["Envisioning the Use of Mobility Data In Practice"](#).

Other teams struggled to find value in the data due to late "kickoff dates" of Facebook mobility data for their specific city or state. This meant that for many teams they did not have access to mobility data at the beginning of the COVID-19 pandemic. This disproportionately affected teams who joined the Network later.

*"We ditched Facebook, pretty much straight off. Just because when we were looking at it, we didn't have that baseline to go from. And so we couldn't compare it across because we didn't want to use these mobility sources by themselves... I did desperately want to use the Facebook data"*

*- Keith Britt, Healthcare Data Analyst, University of Tennessee Medical Center*

*"And so it's not that the data is not useful, but we are missing that before period for comparison. That's one of the limitations."*

*- David Roth, Epidemiologist with Clinical Prevention Services, British Columbia Center for Disease Control*

*"But if the baselines (between different cities or states) were at least the same, you could theoretically compare the outcomes. And unfortunately, the baselines are completely different - they are determined by when they kick off the pipeline for that specific area."*

*- Nishant Kishore, PhD Candidate, Harvard University*

Teams in British Columbia and Tennessee cited this as a major limitation and prompted them to switch or seek other mobility data sources. There are likely more teams that experienced this limitation.

There were likely other factors of non-use which may be due to communication obstacles and data translation among research-practitioner teams. One researcher also cited the negative perception of privacy concerns with Facebook as one of the main reasons they did not pursue using the data in their country. This project was unable to further explore these factors that may have led to non-use.

## MULTIPLE DATA SETS

Some research teams focused on integrating multiple datasets for larger analyses and models. Due to many of the limitations noted above, teams frequently opted for Safegraph and Google data. The research team at the British Columbia Center for Disease Control focused much of their work on creating a pipeline of multiple data sources to help organizations understand movement and behavioral patterns as they pertain to Covid-19 transmission. They used Google mobility trends, municipal transportation data (bus and train), travel data from the Vancouver municipality, and cell phone data. Their Facebook kickoff date was May 2020. Keith Britt from Tennessee used multiple data sources including Google mobility trends and Safegraph and was able to share the number of visits to a specific location as a way to support hospital leadership, the mayor of Knoxville and Knox County in their efforts to identify cluster breakouts and support contact tracing. Keith also briefly explored Unicast as it was also being used by the Governor, but found the data inaccessible due to cost limitation, prompting him to look for other data sources.

*“Our governor was using Unicast. He saw a small blip of mobility. So he put the state on lockdown. So I went to Unicast. And I said, Hey, can I have your data? And they said, Yeah, sure. It'll be \$1000 bucks a week. They didn't want to spend \$52,000 a year on that. So I had to go out and find free sources.”*

*- Keith Britt, Healthcare Data Analyst, University of Tennessee Medical Center.*

## AN EARLY VIEW OF DATA SEEKING BEHAVIORS

Data seeking behaviors appeared to vary between practitioners and researchers. Interviews with practitioners referenced the data seeking behaviors of CIOs and governors specifically for mobility data in what appears to be a boom in a market of mobility data. These types of data seeking behaviors are similar to data scramble activities in the humanitarian information management sector where data minded actors look for various datasets (e.g., public, through private networks) to augment their operational goals in a response. There is less focus on rigorous comparisons between data sources, but a heavy reliance on perceptions of trusted sources of information through organizational and personal networks. It potentially represents a “scramble and see” approach where practitioners may have a more targeted sense of purpose, what may constitute meaningful data, and potentially the actions and decisions that may ensue from using said data. Whether or not this practice of quickly consuming and accepting or rejecting data is advisable is part of a much larger discussion of data use in disasters, as many feel this approach can lead to misguided actions.

The researchers interviewed in this project appeared to spend more time analyzing the quality of the data, probing the metrics, and assessing data veracity. They worked with practitioner teams to varying degrees on achieving value when possible. Their skills in data science, academic inquiry, and the paradigm of *explore, test, verify, publish and act* is a different culture of practice. This type of academic rigor may not be available within data teams in government offices. Data teams in disasters lack the time to investigate using these methods because they have a multitude of operationally focused priorities, which most academics do not have. Few researchers interviewed ingested multiple mobility data sources early in their work in the Network, but focused mostly on one at a time.

*“we brought various mobility and other contact pattern data [into a pipeline] to understand how that affects transmissions, and to get a handle of how behaviors have changed.” - David Roth, Epidemiologist with Clinical Prevention Services, British Columbia Center for Disease Control*

The academic publications about the use of mobility data for COVID-19 from research teams appear to focus on one or only a few mobility data types, diving deep into various metrics, analyses and academic inquiries, rather than the “scramble and research approach”.

More work is needed to explore these data seeking behaviors of both practitioners and researchers as well as that of networks like the CMDN. There are likely a multitude of factors such as familiarity, capacity, skill strengths and limitations, and cultures of practice that likely influence how this broader community seeks data for use in disasters.

*Academia excels at asking questions and coming up with very well reasoned ways to sharpen those questions to a point that they can be answered with the data that we have available. With new data coming in daily, automation of data extraction, cleaning and manipulation are vital pieces of the process which takes research and translates it into production level outputs that can be used by policy makers. Both of these skill sets are necessary to ensure that the best analysis is put in front of the right people in a timely manner.”*

*- Nishant Kishore, PhD Candidate, Harvard University*

**Acknowledgements:** We are grateful to every individual who shared their time amidst the COVID-19 pandemic to provide honest and insightful perspectives. Their voices, enact the change we need to advance translational readiness in future disasters. This project could not have achieved its goals without the support and guidance from Kate Bruni, Lauren Bateman, Abhi Bhatia, Andrew Shroeder, Nishant Kishore, Caroline Buckee, Satchit Balsari, and Alissa Scharf.